Article

Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma

Paraskevaidi, Maria, Ashton, Katherine M, Stringfellow, Helen F, Wood, Nicholas J, Keating, Patrick J, Rowbottom, Anthony W, Martin-Hirsch, Pierre L and Martin, Francis L

Paraskevaidi, Maria, Ashton, Katherine M, Stringfellow, Helen F, Wood, Nicholas J, Keating, Patrick J, Rowbottom, Anthony W, Martin-Hirsch, Pierre L and Martin, Francis L ORCID: 0000-0001-8562-4944 (2018) Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma. Talanta, 189 . pp. 281-288. ISSN 0039-9140

Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma

Maria Paraskevaidi[a,1], Katherine M. Ashton[b], Helen F. Stringfellow[b], Nicholas Wood[c], Patrick Keating[c], Anthony Rowbottom[d], Pierre L. Martin-Hirsch[c] and Francis L. Martin[a,1]

[a]*School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston PR1 2HE, UK*

[b]*Pathology Department, Lancashire Teaching Hospitals NHS Foundation Trust, Preston PR2 9HT, UK*

[c]*Department of Obstetrics and Gynaecology, Lancashire Teaching Hospitals NHS Foundation Trust, Preston PR2 9HT, UK*

[d]*Immunology Laboratory, Pathology Department, Lancashire Teaching Hospitals NHS Foundation Trust, Preston PR2 9HT, UK*

[1]To whom correspondence should be addressed. Email: mparaskevaidi@uclan.ac.uk or flmartin@uclan.ac.uk Tel: +44 (0) 1772 89 6482

Abstract

Robust diagnosis of ovarian cancer is crucial to improve patient outcomes. The lack of a single and accurate diagnostic approach necessitates the advent of novel methods in the field. In the present study, two spectroscopic techniques, Raman and surface-enhanced Raman spectroscopy (SERS) using silver nanoparticles, have been employed to identify signatures linked to cancer in blood. Blood plasma samples were collected from 27 patients with ovarian cancer and 28 with benign gynecological conditions, the majority of which had a prolapse. Early ovarian cancer cases were also included in the cohort (n=17). The derived information was processed to account for differences between cancerous and healthy individuals and a support vector machine (SVM) algorithm was applied for classification. A subgroup analysis using CA-125 levels was also conducted to rule out that the observed segregation was due to CA-125 differences between patients and controls. Both techniques provided satisfactory diagnostic accuracy for the detection of ovarian cancer, with **s**pontaneous Raman achieving 94% sensitivity and 96% specificity and SERS 87% sensitivity and 89% specificity. For early ovarian cancer, Raman achieved sensitivity and specificity of 93% and 97%, respectively, while SERS had 80% sensitivity and 94% specificity. Five spectral biomarkers were detected by both techniques and could be utilised as a panel of markers indicating carcinogenesis. CA-125 levels did not seem to undermine the high classification accuracies. This minimally invasive test may provide an alternative diagnostic and screening tool for ovarian cancer that is superior to other established blood-based biomarkers.

Keywords: Ovarian cancer; diagnostics; biospectroscopy; Raman; SERS

Introduction

Ovarian cancer is frequently discovered at a late stage due to non-specific symptomatology. More than 70% of ovarian cancer patients are diagnosed at an advanced state (stage IV) when the five-year survival rate is 25% [1]. Ideally, disease should be diagnosed promptly and at an early stage (stage I), when cancer is completely confined to the ovaries, as stages II, III and IV are considered advanced with cancer being spread outside the ovaries into the pelvis (*e.g.*, fallopian tubes, bladder or rectum), abdominal cavity (*e.g.*, lining of the abdomen or lymph nodes) and other distinct organs (*e.g.*, lungs), respectively [2]. As a consequence, the five-year survival rate of stage II patients is increased to 70%, while for stage I patients it is further increased to 90% [1, 3].

Screening or diagnostic tests for ovarian cancer comprise of cancer antigen CA-125 alone, ultrasound imaging of the ovaries or a combination. These tests have different screening utility depending on whether they are applied in low or high risk populations. However neither, even in combination, have robust levels of diagnostic accuracy [4, 5]. A variety of blood tests have also been developed with CA 19-9, human epididymis protein 4 (HE4), apolipoprotein A1 (ApoA1), insulin growth factor II (IGF-II) and transferrin being some of them [6-9]. However, most of these individual biomarker tests yield unacceptable diagnostic accuracies which render them unsuitable for clinical use. Recent strategies attempt to combine a number of these tests to achieve superior performance.

Raman spectroscopy has been used extensively in cancer diagnostics utilising a variety of samples (*e.g.*, cells, tissues or biological fluids). Other diagnostic techniques, such as optical coherence tomography, fluorescence microscopy or nonlinear microscopy could also be used for diagnostic purposes, however Raman spectroscopy has been shown in many cases to provide better results [10]. Cervical, skin, breast, oral and brain cancers, as well as other diseases, are some of the wide applications of Raman, facilitating disease detection/monitoring

or even intraoperative assessment of surgical margins [11-19]. Moreover, previous spectroscopic studies have successfully investigated ovarian tissue post-surgery and showed significant differences between normal and malignant cases [20-22]. However, the use of biological fluids, such as blood samples, are numerous: minimally-invasive collection, easier sample preparation and collection of serial samples from the same participants, just to name a few [23, 24]. Raman spectroscopy investigates the phenomenon of inelastic light scattering that is caused after the interaction of light with matter. The sample's electrons first get excited to a virtual state and then fall back to their original energy level either by losing or by gaining energy. The generated shift in energy is characteristic for specific biomolecules such as proteins, lipids and nucleic acids, providing thus invaluable information for a biological sample.

Raman scattering is inherently weak and, therefore, enhancement techniques have been developed to increase the derived signal [25]. Surface-enhanced Raman spectroscopy (SERS) is one of the commonly applied methods which utilises rough metallic surfaces or nanostructures (*e.g.*, silver or gold nanoparticles) to increase the Raman signal by 10^3-10^{10} times. SERS exploits the great electromagnetic field enhancement that is caused by oscillations of surface electrons, called surface plasmons [26]. This allows detection of molecules at low concentration and can partly account for fluorescence that may distort the spectra [15, 27].

The main objective of this study was to use blood spectroscopy in order to assess the diagnostic accuracy for ovarian cancer in women with both early- and late-stage cancer, which has not been previously attempted. Extraction of differential spectral biomarkers was also performed and tentative assignments were made for the development of a panel of diagnostic markers. An important confounding factor, which has not been taken into account in previous spectroscopic studies, and could lead to falsified classification between cancer and healthy controls was the CA-125 level; therefore its effect on the spectral results has been now

calculated in a separate subgroup analysis. To the best of our knowledge this is also the first study employing both Raman and SERS to investigate the effect of the enhanced approach in the diagnostic accuracy – does increased signal necessarily imply improved diagnostic accuracy as well?

Materials and Methods

Population

This study included 27 women with ovarian cancer (17/27 stage I) and 28 women with benign gynecological conditions or a prolapse. All specimens were collected with ethical approval obtained at Royal Preston Hospital UK (16/EE/0010). Mean-age was 68 years for the cancer group and 56 years for the non-cancer group. More information about the cohort characteristics can be found in Table 1; more detailed information about each participant is given in Table S1 [see Supplementary Information (SI)]. Age difference between the different groups was also taken into account to demonstrate whether it affected the spectral results, and therefore the diagnostic accuracy (see SI Fig. S1). Women who were on Tamoxifen have been excluded.

CA-125 measurement

CA-125 levels were determined in blood serum samples for both the ovarian cancer patients and healthy individuals. This test, is a two-site sandwich immunoassay using electrochemiluminescence (ECL) technology which uses monoclonal antibodies (Elecsys CA 125 II, Roche Diagnostics GmbH). The system (Roche Cobas 8000) automatically dispenses 20 µl of sample into a cuvette and then dispenses a biotinylated CA125-specific antibody and a CA125-specific antibody labelled with ruthenium complex react to form a sandwich complex. Streptavidin microparticles are then added and the complex becomes bound to the solid phase *via* the interaction of biotin and streptavidin. The reaction mixture is aspirated in to the reaction

cell where the microparticles are magnetically captured onto the surface of the electrode. Unbound substances are then removed with Procell solution. Application of a voltage to the electrode induces the chemiluminescent emission, which is measured by the photomultiplier; the results are then determined via a calibration curve. A direct relationship exists between the amount of CA-125 in sample and the amount of photons detected by the system. The reference range of CA-125 is 0-35 units/ml (0-35 kU/L), with values >35 kU/L indicating an increased probability for residual or recurrent ovarian cancer in patients treated for primary epithelial ovarian cancer.

Blood plasma preparation for spontaneous Raman and SERS analysis

Whole blood was collected into EDTA tubes, centrifuged at 2000 rpm for 10 min to remove the cells (erythrocytes, white blood cells and platelets) from plasma. The supernatant was then collected and stored at -80^{o}C and thawed at room temperature prior to spectroscopic interrogation. After the samples were thawed, 50 μL were deposited directly on aluminium foil slides and left to air-dry. In order to employ SERS as an enhancement method, silver nanoparticles (AgNPs), with a diameter of 100 nm, were used (nanoComposix, Inc., San Diego). The stock solution (mass concentration: 1.02 mg/ml) was diluted in phosphate buffered saline (PBS); 1 μl AgNPs was diluted in 100 μl PBS. Fifty μl of the diluted solution were mixed with 50 μl of the biological fluid and the resulting mixture (100 μl) was then deposited on aluminium foil slides and was again left to air-dry at room temperature before Raman spectra were acquired.

Spectral acquisition

The experimental settings were kept the same for both Raman and SERS analysis. Spectra were collected with an InVia Renishaw Raman spectrometer coupled with a charge-coupled device (CCD) detector and a Leica microscope. A 785 nm laser was used with a 1200 l/mm grating and the system was calibrated to 520.5 cm^{-1} by using a silicon source before every

run. Seven point spectra were acquired per sample using a 50× objective, 10 second exposure time, 5% laser power and 2 accumulations in the spectral range of 2000-400 cm^{-1} to achieve optimum spectral quality.

Spectral pre-processing and classification

Spectra were evaluated during collection and any cosmic rays were removed by using the Renishaw WiRe software. An in-house developed IRootLab toolbox (http://trevisanj.github.io/irootlab/) was then implemented within MATLAB environment (MathWorks, Natick, USA) for further pre-processing and classification of the data. An initial pre-processing phase is required to deal with any background noise or non-biological variability associated with spectral acquisition or instrumentation. Herein, all spectra were firstly truncated to the biological region (1800-500 cm^{-1}), wavelet denoised, polynomial baseline corrected and vector normalized. All of these steps are standard in the Raman analysis of biological samples in order to generate noise-free spectra with conventional appearance [27]. Difference-between-mean (DBM) spectra was also performed to extract potential biomarkers by subtracting the mean spectra of two classes (*i.e.*, ovarian cancer patients and controls); a peak detection algorithm was implemented to identify the ten most segregating peaks.

Support vector machine (SVM) is a supervised machine-learning technique for creating a classification function from training data. Some of the criteria for the choice of classifier include the achieved diagnostic accuracy, as well as training and computational time [28]. For SVM implementation, the already pre-processed dataset was further normalized (to the [0, 1] range) in order to put all the variables on the same scale. We used the Gaussian kernel SVM, which implies that there are two parameters to be tuned to the value that gives best classification: c and gamma [29]. The optimal tuning parameters were found using grid search (5-fold cross-validation) and then used to calculate the sensitivity and specificity for the different comparisons [29, 30]. Sensitivity is defined as the probability of a test being positive

when the disease is present, while specificity is defined as the probability that a test will be negative at the absence of disease; they were calculated by the following equations:

$$Sensitivity(\%) = \frac{TP}{TP+FN} \times 100 \quad (1)$$

$$Specificity(\%) = \frac{TN}{TN+FP} \times 100 \quad (2)$$

where TP is defined as true positive; FN as false negative; TN as true negative; and FP as false positive.

Statistical analysis

The common peaks that were found to differentiate the classes in both Raman and SERS, after the implementation of the DBM algorithm, were further analyzed in GraphPad Prism 7.0 (GraphPad Software Inc., La Jolla, CA, USA). A student's t-test (non-parametric, two–tailed, 95% confidence interval (CI)) was performed to account for statistical significance with a *P*-value of 0.05 or less being considered significant. Statistical analysis was carried out on averaged spectra in order to account for differences between individuals and not spectra.

Availability of data

All data (raw and pre-processed spectra) along with appropriate code identifiers will be uploaded onto the publicly accessible data repository Figshare.

Results

The enhancement effect of SERS is shown in Fig. 1; after the addition of the AgNPs solution in the blood samples, the Raman signal is notably increased as the silver nanostructures are closely adsorbed to the biomolecules present in the plasma (Fig. 1B). The spectral differences between Raman and SERS spectra were expected and can be attributed to the complex nature of the samples, as well as the nonspecific binding of the nanoparticles to the biomolecules.

196 The classification algorithm was performed in both datasets, spontaneous Raman and
197 SERS, to calculate the sensitivity and specificity rates with which these methods can
198 distinguish ovarian cancer patients (*n*=27) and healthy individuals (*n*=28). For the Raman
199 dataset the achieved sensitivity and specificity were 94% and 96%, respectively (Fig. 2A); for
200 the SERS dataset sensitivity and specificity were 87% and 89%, respectively. After the DBM
201 implementation, ten peaks responsible for the differentiation, were selected; out of those, five
202 peaks were picked up by both and Raman and SERS and, therefore, these were used for further
203 statistical analysis (Fig. 3A and 3B). The five peaks that were selected with Raman
204 spectroscopy were: 1657 cm^{-1} (Amide I, P = 0.0158; 95% CI = 0.00049 to 0.00471), 1418 cm^{-1}
205 (CH$_2$ in lipids, P = 0.0034; 95% CI = 0.00061 to 0.00334), 1301 cm^{-1} (CH$_2$ in lipids, P =
206 0.0612; 95% CI = -0.00379 to 0.00007), 1242 cm^{-1} (Amide III, P = 0.0103; 95% CI = -0.00521
207 to -0.00066) and 916 cm^{-1} (amino acids/carbohydrates, P = 0.0024; 95% CI = 0.00111 to
208 0.0047), while with SERS: 1655 cm^{-1} (Amide I, P = 0.0351; 95% CI = -0.0117 to -0.0005),
209 1429 cm^{-1} (CH$_2$ in lipids, P = 0.066; 95% CI = -0.00049 to 0.00873), 1302 cm^{-1} (CH$_2$ in lipids,
210 P = 0.0882; 95% CI = -0.00825 to 0.00079), 1257 cm^{-1} (Amide III, P = 0.0003; 95% CI = -
211 0.00916 to -0.00283) and 919 cm^{-1} (amino acids/carbohydrates, P = 0.0004; 95% CI = -0.0067
212 to -0.00163).

213 In order to show that the achieved accuracy was not just due to the difference in the
214 CA-125 levels between cancer patients and healthy individuals, we also performed the SVM
215 classification after taking into account the different protein levels. Sensitivity and specificity
216 remained exceptionally high: Raman yielded 99% sensitivity and 85% specificity after
217 comparing individuals with CA-125>35 (Fig. 4A), as well as 78% sensitivity and 99%
218 specificity for individuals with CA-125<35 (Fig. 4B); SERS achieved sensitivity and
219 specificity of 96% and 74%, respectively, for the group with CA-125>35 (Fig. 4C), as well as
220 72% sensitivity and 97% specificity for the CA-125<35 group (Fig. 4D). Similarly, we

considered the age difference between the different groups using the spectra from the spontaneous Raman spectroscopy only (as these provided better results in the previous analyses). The average age of women diagnosed with endometrial cancer is 60 and therefore we considered this as our threshold value. Re-arranging according to age, we had the below cohorts: OC ≥60 years (*n*=20), Control ≥60 years (*n*=10), OC <60 years (*n*=7), Control <60 years (*n*=19). After following the same pre-processing and multivariate analysis as previous, we achieved 98% sensitivity and 90% specificity for the older group (≥60 years) as well as 79% sensitivity and 97% specificity for the younger group (<60 years) (see SI Fig. S1).

Raman and SERS were also used to detect the early ovarian cancer cases (*n*=17) and assess their diagnostic performance. Spontaneous Raman spectroscopy achieved 93% sensitivity and 97% specificity (Fig. 5A), while SERS achieved 80% sensitivity and 94% specificity (Fig. 5B).

Discussion

Although there has been a great effort in developing biomarkers for the early diagnosis of ovarian cancer, there is still no robust method to achieve this. This study has demonstrated the effectiveness of Raman spectroscopic methods toward the diagnosis of ovarian cancer patients, including early cases. Herein, blood plasma samples were used as a minimally invasive way of specimen collection. Blood biospectroscopy, with either infrared (IR) or Raman, has been previously evaluated in gynecological malignancy. Specifically, IR analysis of plasma and serum was applied to diagnose ovarian and endometrial cancers, providing remarkable accuracy (~97% for ovarian and ~82% for endometrial cancer) [31]; SERS analysis of plasma achieved 97% sensitivity and 92% specificity for the segregation of cervical cancer cases from normals [32]; cervical cancer and precancer were also detected with serum sample Raman spectroscopy [33]; both IR and spontaneous Raman were used to analyze blood plasma

and serum towards the diagnosis of ovarian cancer, yielding 93% accuracy for IR spectra and 74% for Raman spectra of plasma [34]; Raman spectroscopy also showed promising results for ovarian cancer diagnosis in 11 patients with the disease, reaching 90% sensitivity and 100% specificity [10]; more recently, it was demonstrated that SERS was able to diagnose endometrial cancer in a pilot study using plasma and serum [35]. Some of the limitations of the above-mentioned studies include either the small number of samples or the absence of a subgroup analysis detecting early stage cases, as well as the lack of CA-125 information as a confounding factor in ovarian cancer. All of these issues have been adequately addressed in the present study. By using a satisfactory number of samples (almost 30 participants in each cohort), we managed to accurately detect both early- and late-stage ovarian cancer cases, which has not been previously shown.

In order to overcome the limitation of low signal in spontaneous Raman, SERS using AgNPs was also employed. Another advantage coming with the use of NPs is that they can be used for more specialised analysis if conjugated with targeting biomolecules, such as antibodies [36]. SERS has been shown to substantially increase the Raman signal and be beneficial for single-molecule detection; however, at the same time it presents with a number of limitations, such as lack of reproducibility and preferential metal-molecule binding, which leads to localised enhancement. This may be the reason for the decreased diagnostic accuracy when compared to spontaneous Raman. The preferential enhancement and lack of repeatability in SERS are also reflected by the increased standard deviation in the class means (Fig. 2B). Sensitivities and specificities were substantially high in both SERS (87% and 89%, respectively) and spontaneous Raman (94% and 96%, respectively), with SERS possibly being more sensitive as a biomarker extraction technique.

Another plausible explanation for the decreased accuracy in SERS is the use of EDTA during plasma collection. EDTA is a molecule for complexing metal ions and it has been found

that its carboxylate groups can bind to nanoparticles surface and be responsible for the generation of new spectral bands [37]. This could potentially obscure the detection of the biological information in the derived spectra. Common anticoagulants used in plasma tubes, such as EDTA and citrate, have been previously found to interfere with SERS spectra; this was suggested to be dealt by the use of serum samples or lithium heparin as the anticoagulant [38].

Blood and its constituents are an invaluable source of information, reflecting alterations in the circulation that can be indicative of a change in health status. Recently, circulating tumour DNA (ctDNA) has attracted much attention as a blood biomarker for early and late stage malignancies, introducing an era of "liquid biopsies" [39, 40]. Also, cell-free DNA (cfDNA), reflecting both normal and ctDNA that is released after cellular necrosis and apoptosis, has been previously found significantly increased in the plasma samples of ovarian cancer patients [41]. A recent systematic review and meta-analysis of nine studies (including 462 ovarian cancer and 407 controls) concluded that cfDNA diagnosed ovarian cancer with 70% sensitivity and 90% specificity and suggested further validation and/or combination with other available biomarkers to improve the diagnostic accuracy [42]. Another study has also shown that ctDNA biomarkers could detect residual tumour, as well as predict response to treatment and survival in ovarian and endometrial cancer cases [43]. With all this in mind, it is quite possible that ctDNA fragments also contributed to the considerably high diagnostic accuracy in this spectroscopic study.

Another scope of the current study was to extract spectral biomarkers, responsible for the differentiation between the malignant and healthy individuals. Each spectral peak corresponds to chemical bonds which are present in specific biomolecules; thus, one can tentatively assign a number of disease biomarkers. To achieve this, the difference between ovarian cancer and control spectra was calculated and the ten most discriminating peaks were selected with a peak-detection algorithm; both Raman and SERS revealed five peaks in

common and these were chosen for further statistical analysis (Fig. 3). The common peaks were correlated to proteins (Amide I and Amide III), lipids and amino acids/carbohydrates. Surprisingly, two out of five spectral regions (~1657-1655 cm^{-1} and ~919-916 cm^{-1}) showed inconsistency between the two spectroscopic approaches; Amide I region was decreased for ovarian cancer patients after Raman spectroscopy, while after SERS the same region was increased. Similarly, the amino acid/carbohydrate region was found decreased in ovarian cancer after Raman and increased after SERS. However, due to the fact that SERS increases significantly the signal of specific peaks, allowing thus more detailed assessment, it is possibly a more sensitive method for biomarker extraction.

More than 160 proteins have been reported to be differentially expressed in ovarian cancer, with some being upregulated, such as CA-125, CA19-9, HE4 or mesothelin, and other being downregulated, such as epidermal growth factor receptor (EGFR) and ApoA1 [44]. Amide I (~1650 cm^{-1}) and Amide III (~1300 cm^{-1}) bands represent protein molecules and are mainly associated with the C=O stretching and C-N stretching/N-H bending vibrations, respectively. The increased level of Amide I and Amide III in ovarian cancer patients after SERS, may correlate with the changes occurring due to the overexpressed proteins. The spectral bands indicative of lipids were both decreased (1429 cm^{-1}) and increased (1302 cm^{-1}) in the ovarian cancer group, which is also backed by previous studies showing a dysregulation of lipid metabolism in cancer [45]. For instance, some studies have shown increased lipid levels in ovarian cancer [45-47], while a limited number of studies have reported reduction [47, 48]. An alternative interpretation of the decreased lipid region (1429 cm^{-1}) could be the downregulation of ApoA1 which has been previously shown to diagnose ovarian cancer in plasma and was estimated at 1484-1427 cm^{-1} [49]. The rise seen in the amino acids/carbohydrate region (919 cm^{-1}) could potentially be attributed to ctDNA, as discussed

previously, or correlated with increased amount of carbohydrates which is considered a risk factor for ovarian cancer [50, 51].

Previous spectroscopic studies investigating ovarian malignancy have not taken into account the differences between CA-125 levels, which may have led to an unrealistic segregation between patients and healthy controls. In order to investigate whether the high diagnostic accuracies achieved in our study were actually attributed to the presence of cancer or just the difference in the CA-125 levels, we also carried out a subgroup analysis to account for this. The extra analysis showed that sensitivities and specificities remained equally satisfactory which denotes that the differences found in our cohort were not attributed to CA-125 but rather to the cancerous condition. Also, after accounting for age differences, it was evident that age alone was not the reason for the high diagnostic accuracy. Even though there is a slight decrease in sensitivity and specificity (*i.e.,* for ≥60 years, specificity dropped from 96% to 90%; for <60 years, sensitivity dropped from 94% to 79%), the diagnostic capability remained very high.

Improved diagnostic performance for the early-stage ovarian cases was a critical objective of this study in order to allow early intervention and potentially improve patient outcomes. Again, both spectroscopic methods provided outstanding diagnostic accuracy, with Raman (sensitivity: 93% and specificity: 97%) being superior to SERS (sensitivity: 80% and specificity: 94%). Current approaches for the early detection of ovarian cancer include biomarker tests, such as serum CA-125 and HE4, imaging techniques, such as computed tomography (CT), transvaginal ultrasound (TVUS) and positron emission tomography (PET) or a combination of these [52]. However, there are still a number of limitations in these methods including expense and lack of optimal sensitivity and specificity. For instance, the sensitivity and specificity of CA-125 is known to be poor, with only 50% of the patients having elevated levels of the protein at stage I and ~75-90% of the cases at a later stage [4]. CA-125 level can be used more reliably to monitor treatment as levels of CA-125 decrease when a treatment is

efficient. However, it is not useful for screening as CA-125 level can be elevated in other conditions, such as endometriosis, breast or lung malignancies, and also not every woman with ovarian cancer has elevated CA-125; CT is expensive and has high false-positive rates which prevent its use in screening [1]. Even though TVUS is preferred than other imaging techniques in terms of speed and sensitivity, there is yet no convincing evidence that it detects early ovarian cancer without causing overtreatment of non-malignant cases [2]. TVUS can indeed show a mass in the ovary but it cannot distinguish whether the mass is benign or malignant. Therefore, other blood biomarkers (CA-125) are used together with ultrasound to identify ovarian tumour at high risk of malignancy. Previous large cohort studies have evaluated the sensitivity and specificity of multimodal screening (MMS) (i.e., annual testing of CA-125 with ultrasound scan as a second line test) and ultrasound screening (USS) (*i.e.*, ultrasound alone); their results showed that the MMS gave slightly higher sensitivity [5, 53]. Specifically, the overall sensitivity for detection of ovarian cancers, diagnosed within a year of a screening, was 84% in the MMS group and 73% in the USS group [5]. However, the positive predictive value for USS was estimated at around ~5%, which indicates a quite high false-positive rate [53]. In an effort to improve the diagnostic accuracy many groups have also combined different biomarkers, which however increase the cost and time requirement [1, 44]. By using spectroscopic techniques these drawbacks seem to be eliminated as they provide a simpler, cost-effective, multi-marker assay, thus securing robustness. The diagnostic accuracy shown in this study is even better than the currently used tests.

In conclusion, the efficacy of Raman spectroscopic methods (*i.e.*, spontaneous Raman and SERS) in detecting ovarian cancer, including early-stage patients, has been demonstrated. Continuous efforts are being made to improve clinical diagnosis and monitoring of disease in ovarian cancer. Our findings suggest improved diagnostic accuracy compared to traditional biomarkers. Specific biomolecules were also found responsible for the segregation between the

cancer and healthy cases and could be used as spectral biomarkers. Future spectroscopic studies should focus on the validation of these results in larger datasets and across different scientific groups and laboratories; this would open a new road in ovarian cancer research and potentially allow the implementation of blood spectroscopy in clinical practice as a promising diagnostic tool.

Acknowledgements

MP acknowledges Rosemere Cancer Foundation for funding. Special thanks to Emma Callery and Debra Beesley for conducting the serum CA-125 tests.

Availability of data and materials

All raw and pre-processed spectra will be available at the publicly accessible data repository Figshare.

Disclosure/Conflict of Interest

The authors declare no conflicts of interest

Authorship

FLM and PLMH conceived the study; MP designed the study, conducted the spectroscopic, multivariate/statistical analysis and wrote the manuscript; KA, HFS, NW, PK, AR and PLMH collected and provided the samples. All authors provided constructive feedback during manuscript preparation. All authors have approved the final version.

References

[1] Z. Yurkovetsky, S. Skates, A. Lomakin, B. Nolen, T. Pulsipher, F. Modugno, J. Marks, A. Godwin, E. Gorelik, I. Jacobs, Development of a multimarker assay for early detection of ovarian cancer, J Clin Oncol 28(13) (2010) 2159-2166.

[2] G.C. Jayson, E.C. Kohn, H.C. Kitchener, J.A. Ledermann, Ovarian cancer, Lancet 384(9951) (2014) 1376-1388.

[3] C.H. Holschneider, J.S. Berek, Ovarian cancer: epidemiology, biology, and prognostic factors, Semin Surg Oncol, Wiley Online Library, 2000, pp. 3-10.

[4] E. Moss, J. Hollingworth, T. Reynolds, The role of CA125 in clinical practice, J. Clin. Pathol. 58(3) (2005) 308-312.

[5] I.J. Jacobs, U. Menon, A. Ryan, A. Gentry-Maharaj, M. Burnell, J.K. Kalsi, N.N. Amso, S. Apostolidou, E. Benjamin, D. Cruickshank, Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial, Lancet 387(10022) (2016) 945-956.

[6] E. Kobayashi, Y. Ueda, S. Matsuzaki, T. Yokoyama, T. Kimura, K. Yoshino, M. Fujita, T. Kimura, T. Enomoto, Biomarkers for screening, diagnosis, and monitoring of ovarian cancer, Cancer Epidemiol Biomarkers Prevent 21(11) (2012) 1902-1912.
[7] F. Su, J. Lang, A. Kumar, C. Ng, B. Hsieh, M.A. Suchard, S.T. Reddy, R. Farias-Eisner, Validation of candidate serum ovarian cancer biomarkers for early detection, Biomark Insights 2 (2007) 369.
[8] V. Nossov, M. Amneus, F. Su, J. Lang, J.M.T. Janco, S.T. Reddy, R. Farias-Eisner, The early detection of ovarian cancer: from traditional methods to proteomics. Can we really do better than serum CA-125?, Am J Obstet Gynecol 199(3) (2008) 215-223.
[9] I. Visintin, Z. Feng, G. Longton, D.C. Ward, A.B. Alvero, Y. Lai, J. Tenthorey, A. Leiser, R. Flores-Saaib, H. Yu, Diagnostic markers for early detection of ovarian cancer, Clin Cancer Res 14(4) (2008) 1065-1072.
[10] I. Ullah, I. Ahmad, H. Nisar, S. Khan, R. Ullah, R. Rashid, H. Mahmood, Computer assisted optical screening of human ovarian cancer using Raman spectroscopy, Photodiagnosis Photodyn Ther 15 (2016) 94-99.
[11] I. Pence, A. Mahadevan-Jansen, Clinical instrumentation and applications of Raman spectroscopy, Chem Soc Rev 45(7) (2016) 1958-1979.
[12] I.P. Santos, E.M. Barroso, T.C. Bakker Schut, P.J. Caspers, C.G.F. van Lanschot, D.H. Choi, M.F. van der Kamp, R.W.H. Smits, R. van Doorn, R.M. Verdijk, V. Noordhoek Hegt, J.H. von der Thusen, C.H.M. van Deurzen, L.B. Koppert, G. van Leenders, P.C. Ewing-Graham, H.C. van Doorn, C.M.F. Dirven, M.B. Busstra, J. Hardillo, A. Sewnaik, I. Ten Hove, H. Mast, D.A. Monserez, C. Meeuwis, T. Nijsten, E.B. Wolvius, R.J. Baatenburg de Jong, G.J. Puppels, S. Koljenovic, Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics, Analyst 142 (2017) 3025-47.
[13] F.L. Cals, T.C. Bakker Schut, J.A. Hardillo, R.J. Baatenburg de Jong, S. Koljenovic, G.J. Puppels, Investigation of the potential of Raman spectroscopy for oral cancer detection in surgical margins, Lab Invest 95(10) (2015) 1186-96.
[14] M.D. Keller, E. Vargis, N. de Matos Granja, R.H. Wilson, M.A. Mycek, M.C. Kelley, A. Mahadevan-Jansen, Development of a spatially offset Raman spectroscopy probe for breast tumor surgical margin evaluation, J Biomed Opt 16(7) (2011) 077006.
[15] K. Kong, C. Kendall, N. Stone, I. Notingher, Raman spectroscopy for medical diagnostics — From in-vitro biofluid assays to in-vivo cancer detection, Adv Drug Del Rev 89 (2015) 121-134.
[16] M. Jermyn, K. Mok, J. Mercier, J. Desroches, J. Pichette, K. Saint-Arnaud, L. Bernstein, M.C. Guiot, K. Petrecca, F. Leblond, Intraoperative brain cancer detection with Raman spectroscopy in humans, Sci Transl Med 7(274) (2015) 274ra19.
[17] F.M. Lyng, D. Traynor, I.R. Ramos, F. Bonnier, H.J. Byrne, Raman spectroscopy for screening and diagnosis of cervical cancer, Anal Bioanal Chem 407(27) (2015) 8279-89.
[18] M. Paraskevaidi, C.L. Morais, D.E. Halliwell, D.M. Mann, D. Allsop, P.L. Martin-Hirsch, F.L. Martin, Raman spectroscopy to diagnose Alzheimer's disease and dementia with Lewy bodies in blood, ACS Chem Neurosci (2018).
[19] M. Paraskevaidi, C. de Morais, K.M.G. de Lima, K. Ashton, H.F. Stringfellow, P.L. Martin-Hirsch, F.L. Martin, Potential of mid-infrared spectroscopy as a non-invasive diagnostic test in urine for endometrial or ovarian cancer, Analyst (2018).
[20] C.M. Krishna, G. Sockalingum, L. Venteo, R.A. Bhat, P. Kushtagi, M. Pluot, M. Manfait, Evaluation of the suitability of ex vivo handled ovarian tissues for optical diagnosis by Raman microspectroscopy, Biopolymers 79(5) (2005) 269-276.
[21] R. Mehrotra, G. Tyagi, D.K. Jangir, R. Dawar, N. Gupta, Analysis of ovarian tumor pathology by Fourier Transform Infrared Spectroscopy, J Ovarian Res 3(1) (2010) 27.
[22] G. Theophilou, K.M.G. Lima, P.L. Martin-Hirsch, H.F. Stringfellow, F.L. Martin, ATR-FTIR spectroscopy coupled with chemometric analysis discriminates normal, borderline and malignant ovarian tissue: classifying subtypes of human cancer, Analyst 141(2) (2016) 585-594.

[23] A.L. Mitchell, K.B. Gajjar, G. Theophilou, F.L. Martin, P.L. Martin-Hirsch, Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting, J Biophotonics 7(3-4) (2014) 153-165.
[24] M. Paraskevaidi, C.L. Morais, K.M. Lima, J.S. Snowden, J.A. Saxon, A.M. Richardson, M. Jones, D.M. Mann, D. Allsop, P.L. Martin-Hirsch, Differential diagnosis of Alzheimer's disease using spectrochemical analysis of blood, Proc Natl Acad Sci USA (2017) 201701517.
[25] N. MacRitchie, G. Grassia, J. Noonan, P. Garside, D. Graham, P. Maffia, Molecular imaging of atherosclerosis: spotlight on Raman spectroscopy and surface-enhanced Raman scattering, Heart (2017) heartjnl-2017-311447.
[26] B. Fazio, C. D'Andrea, A. Foti, E. Messina, A. Irrera, M.G. Donato, V. Villari, N. Micali, O.M. Maragò, P.G. Gucciardi, SERS detection of Biomolecules at Physiological pH via aggregation of Gold Nanorods mediated by Optical Forces and Plasmonic Heating, Sci Rep 6 (2016) 26952.
[27] H.J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N.J. Fullwood, B. Gardner, P.L. Martin-Hirsch, M.J. Walsh, M.R. McAinsh, N. Stone, F.L. Martin, Using Raman spectroscopy to characterize biological materials, Nat Protoc 11(4) (2016) 664-687.
[28] M.J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H.J. Butler, K.M. Dorling, P.R. Fielden, S.W. Fogarty, N.J. Fullwood, K.A. Heys, C. Hughes, P. Lasch, P.L. Martin-Hirsch, B. Obinaju, G.D. Sockalingum, J. Sulé-Suso, R.J. Strong, M.J. Walsh, B.R. Wood, P. Gardner, F.L. Martin, Using Fourier transform IR spectroscopy to analyze biological materials, Nat Protoc 9(8) (2014) 1771-1791.
[29] C.-W. Hsu, C.-C. Chang, C.-J. Lin, A practical guide to support vector classification, (2003), (https://www.researchgate.net/profile/Chenghai_Yang/publication/272039161_Evaluating_unsupervised_and_supervised_image_classification_methods_for_mapping_cotton_root_rot/links/55f2c57408ae0960a3897985/Evaluating-unsupervised-and-supervised-image-classification-methods-for-mapping-cotton-root-rot.pdf).
[30] J. Trevisan, P.P. Angelov, P.L. Carmichael, A.D. Scott, F.L. Martin, Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives, Analyst 137(14) (2012) 3202-15.
[31] K. Gajjar, J. Trevisan, G. Owens, P.J. Keating, N.J. Wood, H.F. Stringfellow, P.L. Martin-Hirsch, F.L. Martin, Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer, Analyst 138(14) (2013) 3917-26.
[32] S. Feng, D. Lin, J. Lin, B. Li, Z. Huang, G. Chen, W. Zhang, L. Wang, J. Pan, R. Chen, Blood plasma surface-enhanced Raman spectroscopy for non-invasive optical detection of cervical cancer, Analyst 138(14) (2013) 3967-3974.
[33] J.L. González-Solís, J.C. Martínez-Espinosa, L.A. Torres-González, A. Aguilar-Lemarroy, L.F. Jave-Suárez, P. Palomares-Anda, Cervical cancer detection based on serum sample Raman spectroscopy, Lasers Med Sci 29(3) (2014) 979-985.
[34] G.L. Owens, K. Gajjar, J. Trevisan, S.W. Fogarty, S.E. Taylor, B. Da Gama-Rose, P.L. Martin-Hirsch, F.L. Martin, Vibrational biospectroscopy coupled with multivariate analysis extracts potentially diagnostic features in blood plasma/serum of ovarian cancer patients, J Biophotonics 7(3-4) (2014) 200-9.
[35] H.J. Butler, S.W. Fogarty, J.G. Kerns, P.L. Martin-Hirsch, N.J. Fullwood, F.L. Martin, Gold nanoparticles as a substrate in bio-analytical near-infrared surface-enhanced Raman spectroscopy, Analyst 140(9) (2015) 3090-3097.
[36] M. Arruebo, M. Valladares, Á. González-Fernández, Antibody-conjugated nanoparticles for biomedical applications, J Nanomater 2009 (2009) 37.
[37] S. Mondal, S. Verma, Catalytic and SERS Activities of Tryptophan-EDTA Capped Silver Nanoparticles, Z Anorg Allg Chem 640(6) (2014) 1095-1101.
[38] A. Bonifacio, S. Dalla Marta, R. Spizzo, S. Cervo, A. Steffan, A. Colombatti, V. Sergo, Surface-enhanced Raman spectroscopy of blood plasma and serum using Ag and Au nanoparticles: a systematic study, Anal Bioanal Chem 406(9-10) (2014) 2355-2365.

[39] C. Bettegowda, M. Sausen, R.J. Leary, I. Kinde, Y. Wang, N. Agrawal, B.R. Bartlett, H. Wang, B. Luber, R.M. Alani, Detection of circulating tumor DNA in early-and late-stage human malignancies, Sci Transl Med 6(224) (2014) 224ra24-224ra24.
[40] L.A. Diaz Jr, A. Bardelli, Liquid biopsies: genotyping circulating tumor DNA, J Clin Oncol 32(6) (2014) 579-586.
[41] A.A. Kamat, M. Baldwin, D. Urbauer, D. Dang, L.Y. Han, A. Godwin, B.Y. Karlan, J.L. Simpson, D.M. Gershenson, R.L. Coleman, Plasma cell-free DNA in ovarian cancer, Cancer 116(8) (2010) 1918-1925.
[42] Q. Zhou, W. Li, B. Leng, W. Zheng, Z. He, M. Zuo, A. Chen, Circulating cell free DNA as the diagnostic marker for ovarian cancer: A systematic review and meta-analysis, PloS one 11(6) (2016) e0155495.
[43] E. Pereira, O. Camacho-Vanegas, S. Anand, R. Sebra, S.C. Camacho, L. Garnar-Wortzel, N. Nair, E. Moshier, M. Wooten, A. Uzilov, Personalized circulating tumor DNA biomarkers dynamically predict treatment response and survival in gynecologic cancers, PLoS One 10(12) (2015) e0145754.
[44] B.M. Nolen, A.E. Lokshin, Protein biomarkers of ovarian cancer: the forest and the trees, Future Oncol 8(1) (2012) 55-71.
[45] C.R. Santos, A. Schulze, Lipid metabolism in cancer, FEBS J. 279(15) (2012) 2610-2623.
[46] I. Delimaris, E. Faviou, G. Antonakos, E. Stathopoulou, A. Zachari, A. Dionyssiou-Asteriou, Oxidized LDL, serum oxidizability and serum lipid levels in patients with breast or ovarian cancer, Clin Biochem 40(15) (2007) 1129-1134.
[47] M. Tania, M. Khan, Y. Song, Association of lipid metabolism with ovarian cancer, Curr Oncol 17(5) (2010) 6.
[48] Q. Hu, M. Wang, M.S. Cho, C. Wang, A.M. Nick, P. Thiagarajan, F.M. Aung, X. Han, A.K. Sood, V. Afshar-Kharghan, Lipid profile of platelets and platelet-derived microparticles in ovarian cancer, BBA clinical 6 (2016) 76-81.
[49] G. Deleris, C. Petibois, Applications of FT-IR spectrometry to plasma contents analysis and monitoring, Vib Spectrosc 32(1) (2003) 129-136.
[50] L. Augustin, J. Polesel, C. Bosetti, C. Kendall, C. La Vecchia, M. Parpinel, E. Conti, M. Montella, S. Franceschi, D. Jenkins, Dietary glycemic index, glycemic load and ovarian cancer risk: a case–control study in Italy, Ann Oncol 14(1) (2003) 78-84.
[51] B. Qin, P.G. Moorman, A.J. Alberg, J.S. Barnholtz-Sloan, M. Bondy, M.L. Cote, E. Funkhouser, E.S. Peters, A.G. Schwartz, P. Terry, Dietary carbohydrate intake, glycaemic load, glycaemic index and ovarian cancer risk in African-American women, Br J Nutr 115(4) (2016) 694-702.
[52] J.V. Jokerst, A.J. Cole, D. Van de Sompel, S.S. Gambhir, Gold nanorods for ovarian cancer detection with photoacoustic imaging and resection guidance via Raman imaging in living mice, ACS Nano 6(11) (2012) 10366-10377.
[53] U. Menon, A. Gentry-Maharaj, R. Hallett, A. Ryan, M. Burnell, A. Sharma, S. Lewis, S. Davies, S. Philpott, A. Lopes, Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), Lancet Oncol 10(4) (2009) 327-340.
[54] Z. Movasaghi, S. Rehman, I.U. Rehman, Raman spectroscopy of biological tissues, Appl Spectrosc Rev 42(5) (2007) 493-541.
[55] S. Li, L. Li, Q. Zeng, Y. Zhang, Z. Guo, Z. Liu, M. Jin, C. Su, L. Lin, J. Xu, S. Liu, Characterization and noninvasive diagnosis of bladder cancer with serum surface enhanced Raman spectroscopy and genetic algorithms, Sci Rep 5 (2015) 9582.
[56] D. Lin, S. Feng, J. Pan, Y. Chen, J. Lin, G. Chen, S. Xie, H. Zeng, R. Chen, Colorectal cancer detection by gold nanoparticle based surface-enhanced Raman spectroscopy of blood serum and statistical analysis, Opt Express 19(14) (2011) 13565-13577.
[57] S. Feng, R. Chen, J. Lin, J. Pan, G. Chen, Y. Li, M. Cheng, Z. Huang, J. Chen, H. Zeng, Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis, Biosens Bioelectron 25(11) (2010) 2414-2419.

[58] J.L. Pichardo-Molina, C. Frausto-Reyes, O. Barbosa-García, R. Huerta-Franco, J.L. González-Trujillo, C.A. Ramírez-Alvarado, G. Gutiérrez-Juárez, C. Medina-Gutiérrez, Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients, Lasers Med Sci 22(4) (2007) 229-236.

Figure Legends

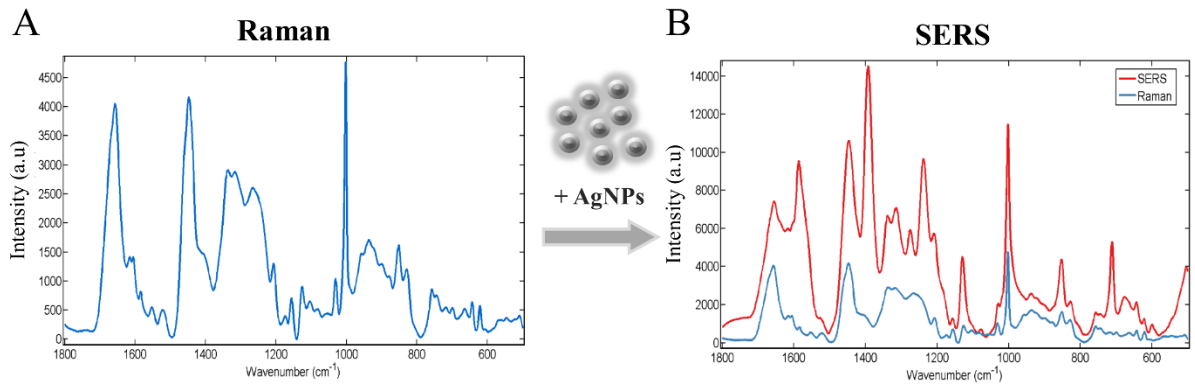
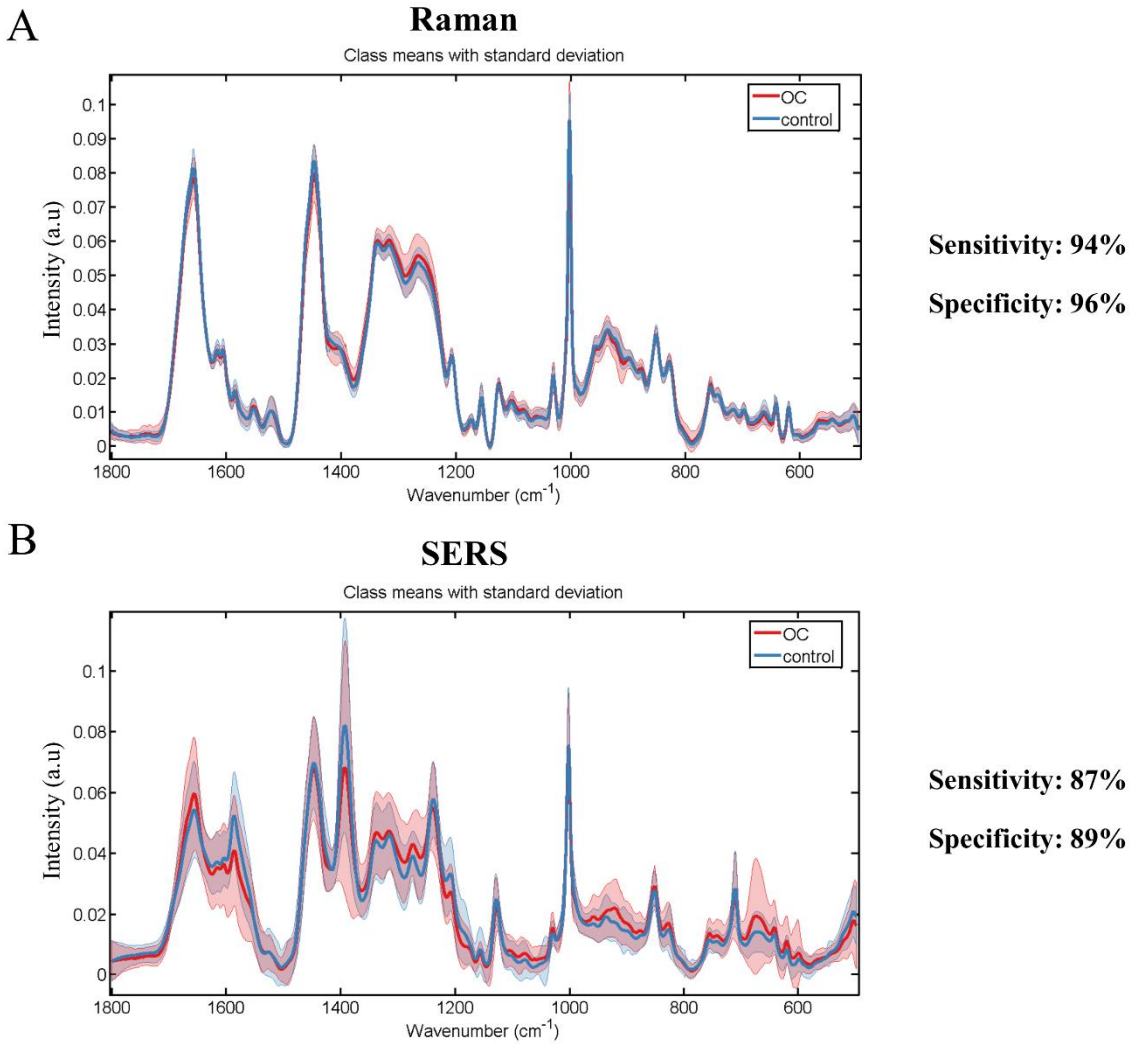


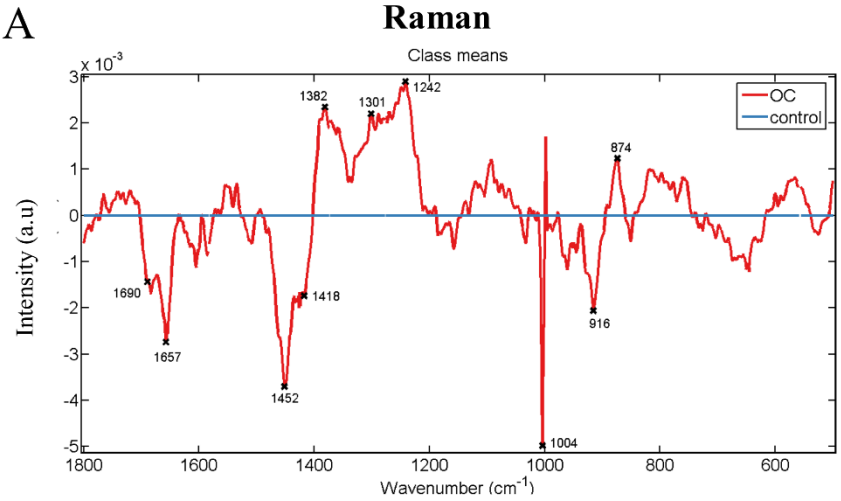
Figure 1: Enhancement effect of SERS after the addition of silver nanoparticles (AgNPs) in blood samples.



591

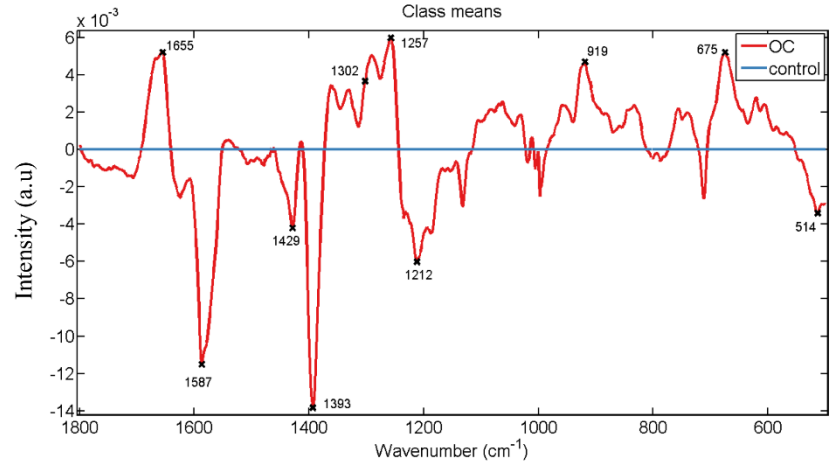
592 **Figure 2:** Diagnostic segregation of ovarian cancer (OC) with (A) Raman spectroscopy and
593 (B) SERS.

A



Wavenumber (cm^{-1})	Tentative Assignment
1690	Amide I/Lipids
1657 *	**Amide I**
1452	$\delta(CH_2)$ (proteins)
1418 **	**CH_2 (lipids)**
1382	$\delta(CH_2)$ (lipids)
1301	**CH_2 (lipids)**
1242 *	**Amide III**
1004	Phenylalanine
916 **	**Amino acids/ Carbohydrates**
874	*ν*(C-C) in amino acids

B



Wavenumber (cm^{-1})	Tentative Assignment
1655 *	**Amide I**
1587	δ(C=C)
1429	**CH_2 (lipids)**
1393	C-H rocking
1302	**CH_2 (lipids)**
1257 ***	**Amide III**
1212	*ν* mode: Tyr/Phe
919 ***	**Amino acids/ Carbohydrates**
675	DNA bases
514	Lipids

Figure 3: Differentiating spectral peaks after (A) Raman spectroscopy and (B) SERS. The tables show the peak positions and tentative assignments of major vibrational bands [54-58]; peaks shown with bold were detected with both Raman and SERS and may be used as more reliable diagnostic biomarkers. Abbreviations: OC: ovarian cancer; *ν*: stretching mode; δ: bending mode.

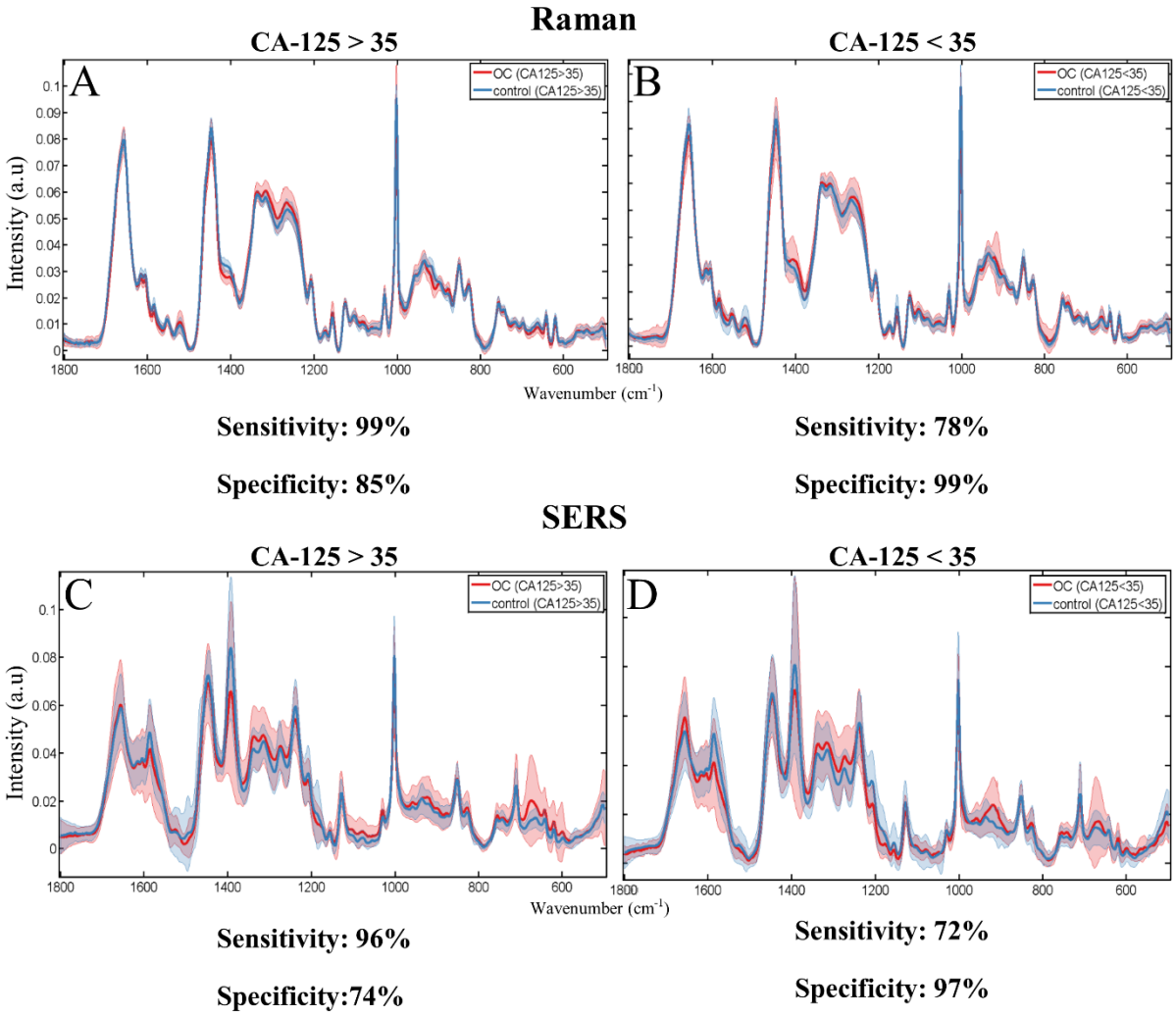
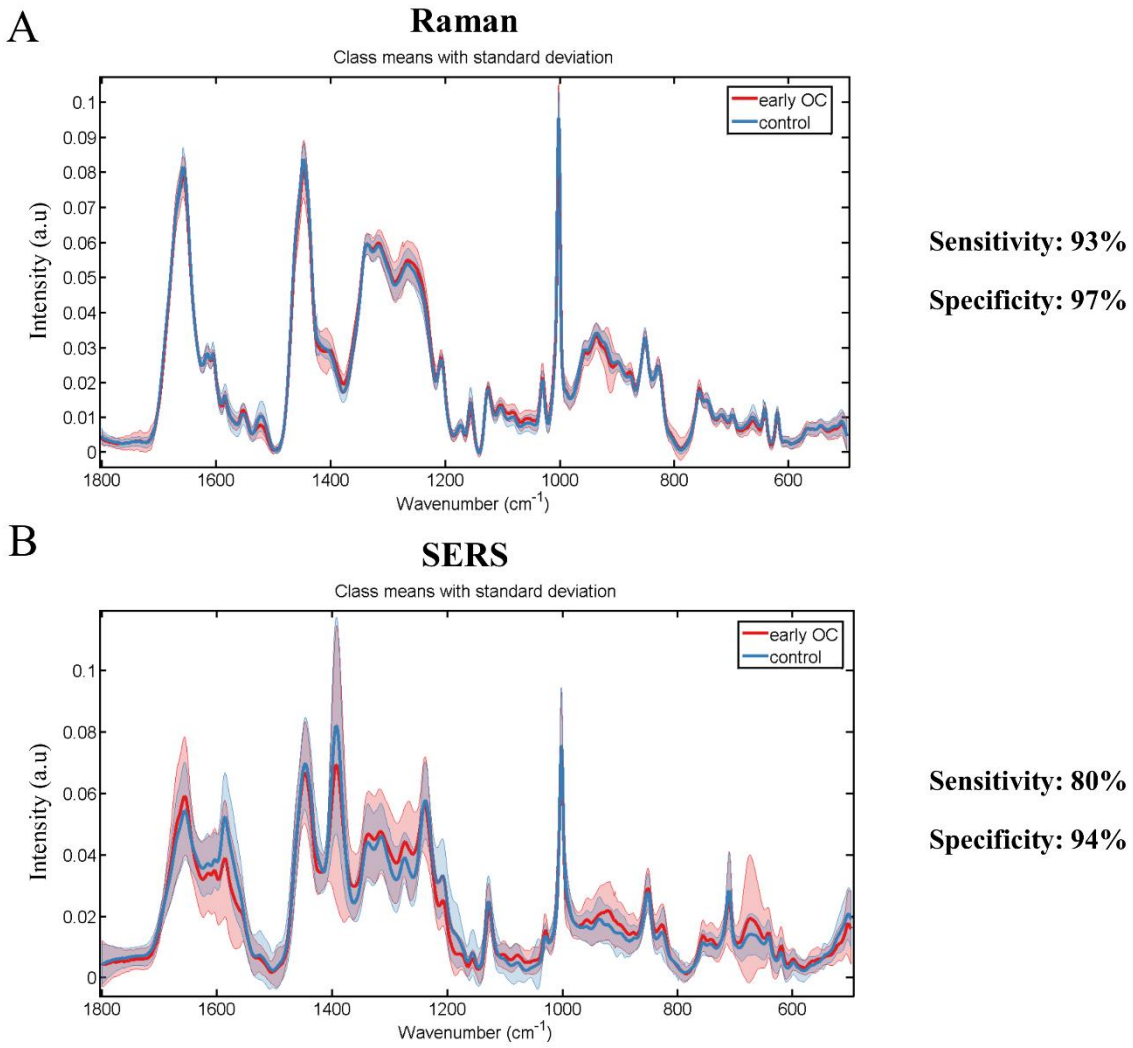


Figure 4: Diagnostic segregation between ovarian cancer (OC) patients and healthy controls according to their CA-125 levels. Sensitivity and specificity are provided for (A) individuals with CA-125>35 u/ml after Raman analysis, (B) individuals with CA-125<35 u/ml after Raman, (C) individuals with CA-125>35 u/ml after SERS and (D) individuals with CA-125<35 u/ml after SERS.



606

607 **Figure 5:** Diagnosis of early ovarian cancer (OC) after (A) Raman spectroscopy and (B) SERS.